

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

## The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation, Mosaic Structure, and Significant Antigenic Variation

By Thomas Proft,\* S. Louise Moffatt,\* Kylie D. Weller,\* A. Paterson,\* Diana Martin,<sup>†</sup> and John D. Fraser\*

From the \*Department of Molecular Medicine, School of Medicine, University of Auckland, Auckland, New Zealand; and the <sup>†</sup>Institute of Environmental Science and Research Limited, Porirua, New Zealand

### Abstract

The frequencies of the newly identified streptococcal superantigen genes *smez*, *spe-g*, and *spe-h* were determined in a panel of 103 clinical isolates collected between 1976 and 1998 at various locations throughout New Zealand. *smez* and *spe-g* were found in every group A *Streptococcus* (GAS) isolate, suggesting a chromosomal location. The *spe-h* gene was found in only 24% of the GAS isolates and is probably located on a mobile DNA element. The *smez* gene displays extensive allelic variation and appears to be in linkage equilibrium with the M/emm type. 22 novel *smez* alleles were identified from 21 different M/emm types in addition to the already reported alleles *smez* and *smez-2* with sequence identities between 94.5 and 99.9%. Three alleles are nonfunctional due to a single base pair deletion. The remaining 21 alleles encode distinct SMEZ variants. The mosaic structure of the *smez* gene suggests that this polymorphism has arisen from homologous recombination events rather than random point mutation. The recently resolved SMEZ-2 crystal structure shows that the polymorphic residues are mainly surface exposed and scattered over the entire protein. The allelic variation did not affect either V $\beta$  specificity or potency, but did result in significant antigenic differences. Neutralizing antibody responses of individual human sera against different SMEZ variants varied significantly. 98% of sera completely neutralized SMEZ-1, but only 85% neutralized SMEZ-2, a very potent variant that has not yet been found in any New Zealand isolate. SMEZ-specific V $\beta$ 8 activity was found in culture supernatants of 66% of the GAS isolates, indicating a potential base for the development of a SMEZ targeting vaccine.

**Key words:** superantigen • streptococcal exotoxin • multiple allelic variation • gene mosaic • antigen variation

### Introduction

Since the 1980s, there has been an alarming increase in group A *Streptococcus* (GAS)<sup>1</sup>-mediated severe invasive disease (SID) around the world (1–5). Streptococcal invasive disease may include soft tissue infections, such as cellulitis, and necrotizing fasciitis or deeper infections, such as septic

scarlet fever (SF) and bacteremia (6). The most severe form of invasive infection results in clinical symptoms similar to those seen in patients with toxic shock syndrome (TSS) caused by *Staphylococcus aureus* and was therefore described as streptococcal toxic shock-like syndrome (STLS) (7, 8). Characteristic symptoms of STLS include hypotension or shock, fever, rash, renal impairment, vomiting, and diarrhea (9). The multiorgan involvement in STLS suggests that a toxin produced by *Streptococcus pyogenes* might be involved in pathogenesis. Prime candidates are the streptococcal pyrogenic exotoxins (SPEs), a family of highly mitogenic proteins secreted individually or in certain combinations by many *S. pyogenes* strains. Together with the staphylococcal enterotoxins (SEs) and the toxic shock syndrome toxin (TSST), they belong to a large family of

Address correspondence to John D. Fraser, Department of Molecular Medicine, School of Medicine, University of Auckland, Private Bag 92010, Auckland, New Zealand. Phone: 64-9-373-7589 ext. 8038; Fax: 64-9-373-7492; E-mail: jd.fraser@auckland.ac.nz

<sup>1</sup>Abbreviations used in this paper: ET, electrophoresis type; GAS, group A *Streptococcus*; GST, glutathione S-transferase; KD, Kawasaki disease; MNT, M-nontypable; SAg, superantigen; SE, staphylococcal enterotoxin; SF, scarlet fever; SID, severe invasive disease; SMEZ, streptococcal mitogenic exotoxin; SPE, streptococcal pyrogenic exotoxin; STLS, streptococcal toxic shock-like syndrome; TSS, toxic shock syndrome.

proteins, also known as bacterial superantigens (SAGs), with similar amino acid sequence, protein structure, and functional activities (10, 11). SAGs simultaneously bind to MHC class II antigens and TCR molecules and thereby stimulate a large number of T cells in a TCR V $\beta$ -specific mode. This event leads to the release of high systemic levels of cytokines, such as TNF- $\alpha$  and IL-1 $\beta$ , and of T cell mediators like IL-2 and IFN- $\gamma$  (12-14).

It has also been suggested that bacterial SAGs are implicated in other diseases, such as acute rheumatic fever and Kawasaki disease (KD) (6, 15, 16). SPE-A has long been associated with the ability of GAS to cause SF and STLS, and is believed to be the primary cause of the skin rash in SF (8, 17, 18). Although it has been shown that SPEs can cause pathological conditions including fever and shock (19) and that severe streptococcal infections are associated with certain exotoxin-producing strains (17, 20, 21), a correlation between disease and a particular exotoxin and/or *S. pyogenes* strain remains elusive. Yu and Ferratti performed molecular epidemiologic analysis on 300 general *S. pyogenes* clinical isolates (from patients with a variety of diseases, except SF) and found the *speA* gene with a frequency of only 15% (21). Among 146 isolates from patients with SF, 45% were shown to contain *speA*. Hauser et al. analyzed 34 *S. pyogenes* isolates from patients with STLS and found the *speA* gene in 85%. The *speC* gene was found in only 21% of the isolates (20). Several investigators reported a correlation between STLS and serotypes M1, M3, T3, and M12 and that *speA* was mostly, but not exclusively, found in M1- and M3-type strains from patients with SF (17, 20, 21). On the other hand, a genetic and phenotypic diversity among *S. pyogenes* isolates from severe invasive infections and STLS was reported in the United States (22), and Forni et al. found an M-nontypable (MNT) strain to be common in STLS patients (9). Furthermore, Hsueh and co-workers found no association of severe invasive disease in Taiwan with any SPE (23).

Four allelic variants of the *speA* gene have been found in *S. pyogenes* so far (24). The *speA1* allele occurs in many distinct clonal lines, but expression of SPE-A2 and SPE-A3, which are characterized by single amino acid substitutions, was shown in two single clones that have caused the majority of STLS episodes in the last decade (24, 25). Musser et al. reported that all contemporary strains assigned to electrophoresis types (ETs) 1 and 2 (two highly pathogenic clones, which generally express serotype M1 and M3, respectively) have the *speA2* and *speA3* allelic variants, respectively (25). In contrast, ET1 and 2 isolates from disease episodes in the 1920s and 1930s contain the *speA1* allele, suggesting temporal variation. Kline and Collins showed that SPE-A3 has a significantly increased mitogenic activity and MHC class II affinity compared with SPE-A1 and SPE-A2 (26). These results suggest that SPE-A3 represents a more active form of SPE-A and might explain the fluctuation in severity and disease frequency of severe invasive infections.

The implication of other exotoxins, like SPE-C and SSA, in streptococcal infections is even less clear and is as-

sumed on the basis that not all cultures from patients with STLS or other severe invasive diseases express SPE-A.

Since some cases of SID cannot be correlated to any of the known SPEs, there is a suggestion that one or more novel SPEs might play a role in the resurgence of SID in recent years. Recently, we reported the identification of the streptococcal exotoxins SMEZ-2, SPE-G, SPE-H, and SPE-J (27). SMEZ-2 is an allelic variant of the previously described SMEZ (28). Functional analysis using recombinant protein revealed that SPE-G and SPE-H target human T cells carrying the V $\beta$ 2 region, a specificity they share with SPE-C and TSST. SMEZ and SMEZ-2 target human V $\beta$ 4 and V $\beta$ 8 T cells, a unique specificity among streptococcal SAGs. A selective expansion of V $\beta$ 2 and V $\beta$ 8 T cells has been reported in patients with KD (15), but the potential role of SAGs in triggering this disease remains unclear.

In this report, we present evidence for wide genotypic variation in SMEZ among a panel of 103 clinical isolates selected to provide a spectrum of distinct M/emm types occurring in the community and to provide representation of subtypes within specific M types. Thus far, 24 *smez* alleles have been identified in 22 different M/emm types. Moreover, the frequency of *smez* in 95 analyzed GAS isolates approached 100%, making it the most common of all known *sag* genes, at least within New Zealand. Given the fact that SMEZ is also the most potent SAG produced by *S. pyogenes*, this has significant implications for its involvement in *Streptococcus*-mediated disease.

## Materials and Methods

**Bacterial Isolates.** The 103 streptococcal isolates were obtained from the Institute of Environmental Science and Research in Porirua, New Zealand. Group 1 consists of 52 isolates (including 6  $\beta$ -hemolytic non-GAS strains) that were taken from various parts of New Zealand between 1976 and 1996. Group 2 contains 51 isolates (including 4 non-GAS) collected from South Auckland schools in 1998. The non-GAS were *Streptococcus agalactiae* (group B), *Streptococcus equis* (group C), and *Streptococcus spp.* (group G). Genotyping of all isolates was undertaken blinded to their strain definition. Two duplicates were included to ensure accuracy of the testing. The results of the duplicate isolates were not included in the overall statistics.

**Genomic DNA Preparation.** All *S. pyogenes* isolates were grown overnight in 10 ml of Brain Heart Infusion medium (BHI; Difco) at 37°C in 15-ml Falcon tubes without agitation. The cells were pelleted and washed in 1 ml of lysis buffer (10 mM Tris, pH 8, 50 mM EDTA) in a 1.5-ml Eppendorf tube and finally resuspended in 400  $\mu$ l lysis buffer, 4 U of mutanolysin (Sigma-Aldrich), 2 mg of lysozyme, and 40  $\mu$ g of RNaseA (Sigma-Aldrich) were added. After incubating the samples for 1 h at 37°C with gentle agitation, 20  $\mu$ g of proteinase K (Sigma-Aldrich) was added and incubated for an additional 30 min. Then, 40  $\mu$ l of a 20% sarcosyl solution was added, and the DNA was extracted with phenol/chloroform and finally precipitated with ethanol. The DNA was resuspended in 50  $\mu$ l of TE buffer.

**Genotyping of *S. pyogenes* Isolates.** Purified genomic DNA from all *S. pyogenes* isolates was used for PCR with specific primers for the *smez*, *spe-g*, and *spe-h* genes as described previously (27). In addition, a primer pair specific to a DNA region encod-

ing the 23S rRNA. oligo 23rRNA fw (GCTATTTCCGA-CAGAACCCAG) and oligo 23rRNA rev (CTGAAACAT-CTAACTACCTG), was designed and used for PCR as a positive control.

**Southern Blot Analysis.** About 5 µg of genomic DNA was digested using restriction enzyme HindIII (GIBCO BRL) and loaded onto a 0.7% agarose gel. The DNA was transferred from the gel to a Hybond-N+ nylon membrane (Amersham Pharmacia Biotech) as described by Maniatis et al. (29). A 640-bp DNA fragment of the *smex-2* gene was radiolabeled using the RadPrime Labeling System (GIBCO BRL) and [ $\alpha$ - $^{32}$ P]dCTP (NEN Life Science Products). The nylon blots were hybridized with the radiolabeled probe in 2× SSC, 0.5% SDS, 5× Denhardt's overnight at 65°C. After washing twice in 0.2× SSC, 0.1% SDS at 65°C, the blots were analyzed on a Storm Phosphor-Imager<sup>®</sup> (Molecular Dynamics).

**Cloning and Sequencing of *smex* Alleles.** All PCR products obtained from genotyping with different *S. pyogenes* isolates were enriched using the Wizard PCR DNA purification system (Promega) and cloned into T-tailed pBlueScript SKII vectors (Stratagene) followed by transformation into *Escherichia coli* XL1 blue.

The *smex* genes were analyzed by the dideoxy chain termination method using a Licor automated DNA sequencer (model 4200). Three subclones of every cloning experiment were analyzed to insure that no Taq polymerase-related mutations were introduced. The DNA sequences of the *smex* alleles have been annotated in EMBL/GenBank/DBJ. The accession numbers are listed in Table II.

**Expression and Purification of Recombinant SMEZ Proteins.** Recombinant toxins were produced as described previously (27). In brief, individual *smex* genes were cloned from pBlueScript vector into the BamHI-EcoRI cloning site of pGEX-2T expression vector and transformed into *E. coli* DH5 cells. For expression of glutathione *S*-transferase (GST) fusion proteins, the DH5 cells were grown in tryptone phosphate medium at 30°C and induced with 0.1 mM isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) at OD<sub>600</sub> of 0.7–0.8. The GST fusion proteins were purified on glutathione (GSH) agarose. The mature proteins were then cleaved off from GST by trypsin digestion and further purified by two rounds of cation exchange chromatography. Due to their low isoelectric point, some SMEZ alleles could not be separated from GST by this method. In these cases, trypsin cleavage was performed on the GSH agarose and the mature toxins were eluted from the column.

**Toxin Proliferation Assay.** Human PBLs were purified from blood of a healthy donor by Histopaque Ficoll (Sigma-Aldrich) fractionation. The PBLs were incubated in 96-well round-bottomed microtiter plates at 10<sup>5</sup> cells per well with RPMI-10 (RPMI with 10% FCS) containing varying dilutions of recombinant toxins. The dilution series was performed in 1:5 steps from a starting concentration of 10 ng/ml of toxin. After 3 d, 0.1 µCi [ $^3$ H]thymidine was added to each well and cells were incubated for an additional 24 h. Cells were harvested and counted on a scintillation counter.

**Jurkat Cell Assay.** Jurkat cells (a human T cell line) and LG-2 cells (a human B lymphoblastoid cell line, homozygous for HLA-DR1) were harvested in log phase and resuspended in RPMI-10. 100 µl of the cell suspension, containing 1 × 10<sup>5</sup> Jurkat cells and 2 × 10<sup>4</sup> LG-2 cells, was mixed with 100 µl of varying dilutions of recombinant toxins on 96-well plates. After incubating overnight at 37°C, 100 µl aliquots were transferred onto a fresh plate and 100 µl (1 × 10<sup>4</sup>) of SeI cells (IL-2-dependent murine T cell

line) per well was added. After incubating for 24 h, 0.1 µCi [ $^3$ H]thymidine was added to each well and cells were incubated for an additional 24 h. Cells were harvested and counted on a scintillation counter. As a control, a dilution series of IL-2 was incubated with SeI cells.

**For detection of native Vβ8-specific exotoxin expression in *S. pyogenes*, each isolate was grown in 10 ml of BHI medium overnight at 37°C. Cells were then removed by centrifugation and supernatants were diluted 1:10, 1:50, 1:250, and 1:1,000 in RPMI-10, and the assay was continued as described above.**

**Neutralization Assay.** PBLs were obtained and stimulated as described under Toxin Proliferation Assay above, with the exception that the 10% FCS was replaced by 10% human serum from various healthy donors. All stimulations were done using 40 pg/ml of recombinant toxin. PBLs from a single donor were used for all tests. The neutralizing response was determined by comparing the T cell proliferation with a control test using 10% FCS instead of 10% human serum. The mitogenic response obtained in the absence of human serum was set to 100%.

## Results

**Genotyping of *S. pyogenes* Isolates.** PCR-based genotyping was performed in order to determine the frequency of the recently identified *sag* genes *smex*, *spe-g*, and *spe-h* in streptococcal isolates (Table I). The PCR primers for *smex* were designed to anneal with both genes, *smex* and *smex-2*. The 103 isolates were collected in New Zealand between 1976 and 1998 from varying sites in patients with varying infections, although the majority were from sore throats. They comprised 94 GAS and 9 non-GAS, which were *S. agalactiae* (group B), *S. equi* (group C), and *Streptococcus* spp. (group G). There are 22 distinct M/emm types represented among the GAS isolates, 13 isolates are MNT, and in 2 cases the M type is unknown. The analysis was undertaken blinded to the details of each isolate, and two duplicate isolates were included (95/31 and 4202) to demonstrate the reproducibility of the testing procedure. The isolates are listed in two groups. Group 1 contained isolates collected within a large time frame (1976–1998) from all over New Zealand. Group 2 comprised isolates collected within a short time (1998) in a distinct area (South Auckland schools).

All of the nine non-GAS isolates (belonging to groups B, C, and G) were negative for the tested *sag* genes. The frequencies for *smex*, *spe-g*, and *spe-h* within the GAS isolates were 95.6, 100, and 23.9%, respectively. A correlation between a certain M/emm type and the presence of the *spe-h* gene could not be established. The deficiency in this current set was that only five M/emm types were represented by more than one isolate. The most frequent serotype was M/emm12 with 13 isolates, from which 7 were positive and 6 were negative for *spe-h*, suggesting genetic diversity within the M/emm12 strain. In contrast, all 12 tested NZ1437/M89 isolates were negative for *spe-h*.

The high frequencies of *smex* and *spe-g* are of particular interest, as this has not been described for any other streptococcal *sag* gene thus far. Other *spe* genes, like *speA*, *speC*, and *ssa*, are found at much lower frequencies, and horizon-

Table 1. Genotyping of *Streptococcus* Isolates

Group 1: Isolates collected from various parts of New Zealand between 1976 and 1996												Group 2: Isolates collected from South Auckland schools in 1998											
Strain no.	Group	M/ strain	Site	Disease	Rib. DNA	spe-g	spe-h	spe-i	V88	Gene		Strain no.	Student ID	Group	M/ strain	Site	Disease	Rib. DNA	spe-g	spe-h	spe-i	V88	Gene
PP 1943	A	M53	u	ST	+	+	-	+	-	n.d.		4202*	3310	A	NZ3118'	u	ST	+	+	-	+	+	spe-14
PP 2658	A	M59	u	ST	+	+	-	+	-	spe-12		4202(2)*	3310	A	NZ3118'	u	ST	+	+	-	+	+	spe-14
FP 4223	A	M80	u	ST	+	+	-	+	-	n.d.		9608	2252	A	MNT	u	ST	+	+	-	+	-	spe-6
FP 5417	A	M41	u	ST	-	-	-	+	+	spe-13		9639	2184	A	MNT	u	ST	+	+	+	+	+	spe-3
FP 6847	A	M1	u	ST	-	-	-	+	+	n.d.		9778	3230	A	unm56	u	ST	+	+	-	+	+	spe-11
FP 6971	A	M57	u	ST	+	+	+	+	+	n.d.		9893	0144	A	PT180	u	ST	+	+	+	+	+	spe-4
1/6045	A	M4	u	ST	+	+	-	+	+	spe-8		9894	6564	A	cmn59	u	ST	+	+	-	+	+	n.d.
79/1575	A	M1	u	Tonsillitis	+	+	+	+	+	spe-1		10019	6264	A	cmn44	u	ST	+	+	+	+	-	spe-19
81/2033	A	M12	u	ST	+	+	+	+	+	n.d.		10028	9386	A	cmn41	u	ST	+	+	-	+	+	n.d.
82/20	A	M4	u	Ulcer	+	+	-	+	+	n.d.		10134	1880	A	ST4547	u	ST	+	+	-	+	+	spe-5
82/332	A	M12	u	ST	+	+	+	+	+	n.d.		10303	3564	A	cmn59	u	ST	+	+	-	+	+	spe-12
82/675	A	NZ1437*	u	Wound	+	+	-	+	+	n.d.		10307	4850	A	NZ3118'	u	ST	+	+	-	+	+	n.d.
84/141	A	M12	u	ST	-	+	+	+	+	n.d.		10438	4904	A	ST3015	u	ST	+	+	-	+	+	spe-24
84/1733	A	M4	u	ST	+	+	-	+	+	n.d.		10463	TSP	A	cmn48	u	ST	+	+	-	+	+	spe-22
84/781	A	NZ1437*	u	ST	+	+	-	+	+	n.d.		10648	11510	A	ST2667	u	ST	+	+	-	+	+	spe-16
85/1	A	M12	u	ST	+	+	-	+	+	n.d.		10730	11503	A	MNT	u	ST	+	+	-	+	+	n.d.
85/167	A	M12	u	ST	+	+	+	+	+	n.d.		10742	2374	A	ST609	u	ST	+	+	-	+	+	n.d.
85/314	A	NZ1437*	u	Wound	+	+	-	+	+	n.d.		10761	3254	A	MNT	u	ST	+	+	-	+	+	spe-6
85/437	A	M81	u	Intr. oesophagus	+	+	-	+	+	spe-9		10763	6614	A	PT3876	u	ST	+	+	-	+	+	spe-18
85/722	A	n.d.	?	?	+	+	-	+	+	spe-23		10782	4890	A	MNT	u	ST	+	+	+	+	+	n.d.
86/435	A	M4	u	RF	-	-	-	+	+	spe-6		10781	10290	A	MNT	u	ST	+	+	+	+	+	n.d.
87/169	A	M12	u	ST	+	+	+	+	+	spe-3		10782	10303	A	MNT	u	ST	+	+	+	+	+	spe-6
87/19	A	M12	u	ST	+	+	+	+	+	spe-3		10848	8854	A	NZ1437*	u	ST	+	+	-	+	+	n.d.
87/781	A	M12	u	ST	+	+	-	+	+	n.d.		10802	6264	A	NZ3118'	u	ST	+	+	-	+	+	n.d.
88/627	A	M12	u	Wound	+	+	-	+	+	n.d.		10889	5194	A	PT2841	u	ST	+	+	-	+	+	spe-20
89/82	A	M12	u	Pneum.	+	+	-	+	+	n.d.		11070	1434	A	cmn53	u	ST	+	+	+	-	-	n.d.
89/85	A	M12	u	Erysipelas	+	+	+	+	+	n.d.		11072	1880	A	ST4547	u	ST	+	+	-	+	+	spe-5
89/86	A	M1	u	ACN	+	+	-	+	+	n.d.		11083	4358	A	MNT	u	ST	+	+	-	+	+	n.d.
89/84	A	NZ1437*	u	ST	+	+	-	+	+	n.d.		11093	9781	A	MNT	u	ST	+	+	+	+	+	n.d.
90/306	A	M3	u	Otitis media	+	+	-	+	+	spe-4		11162	8020	A	PT2612	u	ST	+	+	+	-	-	n.d.
90/424	A	M4	u	ST	+	+	-	+	+	n.d.		11222	4928	A	NZ3118'	u	ST	+	+	+	+	+	n.d.
91/542	A	M12	u	ST	+	+	-	+	+	n.d.		11227	8854	A	cmn14	u	ST	+	+	-	+	+	spe-21
94/11	A	NZ1437*	u	Abcess	+	+	+	+	+	n.d.		11244	2252	A	ST4547	u	ST	+	+	-	+	+	spe-5
94/228	A	M49	u	Endometrit.	+	+	-	+	+	n.d.		11276	4624	A	MNT	u	ST	+	+	-	+	+	n.d.
94/330	A	M4	u	ST	+	+	-	+	+	n.d.		11259	2950	A	unm80	u	ST	+	+	-	+	+	spe-9
94/354	A	M12	u	ST	+	+	-	+	+	n.d.		11674	3186	A	ST609	u	ST	+	+	-	+	+	spe-7
94/384	A	M4	u	Wound	+	+	-	+	+	n.d.		11580	3280	A	cmn53	u	ST	+	+	+	-	-	spe-10
94/712	A	NZ1437*	u	Cellulitis	+	+	-	+	+	n.d.		11610	2424	A	cmn57	u	ST	+	+	-	+	+	n.d.
95/127	A	NZ1437*	u	Cellulitis	+	+	-	+	+	n.d.		11646	1880	A	ST4547	u	ST	+	+	-	+	+	n.d.
95/311	A	NZ1437*	u	Abcess	+	+	-	+	+	spe-13		11681	3664	A	cmn12	u	ST	+	+	-	+	+	spe-3
95/311(2)	A	NZ1437*	u	Abcess	+	+	-	+	+	n.d.		11688	8328	A	PT5757	u	ST	+	+	-	+	+	spe-17
95/381	A	NZ1437*	u	Abcess	+	+	-	+	+	n.d.		11745	12387	A	unm59	u	ST	+	+	-	+	+	n.d.
96/1	A	n.d.	?	?	+	+	-	+	+	n.d.		11789	1568	A	MNT	u	ST	+	+	-	+	+	spe-6
96/384	A	NZ1437*	u	Burns	+	+	-	+	+	n.d.		11802	3266	A	MNT	u	ST	+	+	-	+	+	n.d.
96/881	A	M4	u	Eye infect.	+	+	-	+	+	n.d.		11869	2950	A	ST4547	u	ST	+	+	-	+	+	spe-5
96/810	A	M4	u	SF	+	+	-	+	+	n.d.		11881	4916	A	MNT	u	ST	+	+	-	+	+	n.d.
D21	A	M1	u	Tonsillitis	+	+	-	+	+	n.d.		12016	12373	A	cmn59	u	ST	+	+	+	+	-	n.d.
RC4003	C	-	u	ST	+	-	-	-	-	-		7625	8215	B	-	u	ST	+	-	-	-	-	-
SP2025	C	-	u	ST	+	-	-	-	-	-		8011	3238	B	-	u	ST	+	-	-	-	-	-
N16174	C	-	u	ST	+	-	-	-	-	-		10368	1663	C	-	u	ST	+	-	-	-	-	-
N16182	B	-	u	ST	+	-	-	-	-	-		012632	4395	B	-	u	ST	+	-	-	-	-	-
VC4141	C	-	u	ST	+	-	-	-	-	-													

The isolates were collected in New Zealand between 1976 and 1996 (group 1) and from South Auckland schools in 1998 (group 2) from patients with varying diseases. The results are based on PCR analysis using purified genomic DNA and specific primers for each of the *spe* genes. The non-GAS are: B, *S. agalactiae*; C, *S. equis*; C, *Streptococcus* spp. Site and disease abbreviations are as follows: u, throat site; w, wound site; sk, skin; ps, pus site; hvs, high vaginal site; hc, blood culture; ST, sore throat; RF, rheumatic fever; AGN, acute glomerulonephritis; T, tonsillitis; throat carriage. n.d., not determined.

\*Duplicate isolates.

†Recently assigned as M92.

‡Recently assigned as M89.

§Duplicate isolates.

tal gene transfer might explain the varying frequencies of these genes in different strains. In contrast, both *smez* and *spe-g* were found in virtually all tested GAS isolates. Only four GAS isolates (11152, 11070, 94/229, and 11610) tested negative for *smez*. These were PT2612, emm65, M49, and emm57. Southern hybridization was performed to find out if the negative PCR results were due to lack of the *smez* gene or to lack/alteration of the primer binding site(s). HindIII-digested genomic DNA of selected streptococcal isolates was probed with a 640-bp radiolabeled *smez*-2 PCR fragment (Fig. 1). The *smez* gene is located on a 1953-bp HindIII DNA fragment in the M1 strain (700294; American Type Culture Collection) corresponding to fragment A in the Southern blot. In all four PCR-negative isolates, the *smez* probe hybridized to a HindIII fragment of ~4 kb (fragment B), but not to the *smez* bearing fragment A (samples 17, 26, 57, and 67). In addition, the *smez* probe bound to a second DNA fragment of ~4.2 kb (fragment C) in isolate 11152 (sample 17). In the M1 reference strain (lane 1) and in isolate 4202 (sample 50), the *smez* probe also bound to fragment B. In addition to fragment A. Fragment B in the M1 strain contains a 180-bp region that shares 97% sequence homology with the 3' end of the *smez* gene. These results suggest that the four PCR-negative iso-

lates possess a truncated *smez* gene or a *smez*-like sequence, but not a complete *smez* gene.

**Identification of Multiple *smez* Alleles.** DNA sequence analysis of the *smez* genes from different *S. pyogenes* isolates was performed. 37 isolates were analyzed, representing 22 M/emm types, 5 MNT isolates, and 2 isolates of unknown M type (Table II). 23 different alleles were identified, bringing the total number of *smez* alleles to 24 (*smez* is hereafter referred to as *smez*-1). The *smez* gene appears to be in linkage equilibrium with the M/emm type. Only three *smez* alleles (*smez*-3, *smez*-4, and *smez*-8) were found in more than one M/emm type. *smez*-3 was found in all three tested M/emm12 type isolates plus one MNT. *smez*-4 was found in one M5 and in one PT180. *smez*-8 was found

Table II. List of All *smez* Variants Identified To Date

Gene	Protein	pI	P <sub>50</sub>	M/emm type	Accession no.
<i>smez</i> -1	SMEZ-1	5.4	0.1	M1	- <sup>1</sup>
<i>smez</i> -2	SMEZ-2	6.98	0.02*	n.d.	AF086626
				MNT,	
<i>smez</i> -3	SMEZ-3	6.98	0.25	emm12, M12	AF143653
<i>smez</i> -4	SMEZ-4	5.73	n.d.	PT180, M5	AF143654
<i>smez</i> -5	SMEZ-5	5.17	0.7	ST4547	AF143655
<i>smez</i> -6	n.f.	-	-	MNT	AF143656
<i>smez</i> -7	SMEZ-7	6.14	0.3	ST809	AF143657
<i>smez</i> -8	SMEZ-8	5.17	0.25	M4, M81	AF143658
<i>smez</i> -9	SMEZ-9	8.86	0.2	emm80, M80	AF143659
<i>smez</i> -10	SMEZ-10	5.4	n.d.	emm53, M53	AF143660
<i>smez</i> -11	SMEZ-11	8.86	n.d.	emm56	AF143661
<i>smez</i> -12	SMEZ-12	5.17	n.d.	emm59, M59	AF143662
<i>smez</i> -13	SMEZ-13	6.14	0.2	NZ1437	AF143663
<i>smez</i> -14	SMEZ-14	6.14	n.d.	NZ5118	AF143664
<i>smez</i> -15	SMEZ-15	5.02	n.d.	M41	AF143665
<i>smez</i> -16	SMEZ-16	7.82	n.d.	ST2267	AF143666
<i>smez</i> -17	SMEZ-17	8.86	n.d.	PT5757	AF143667
<i>smez</i> -18	SMEZ-18	9.62	n.d.	PT3875	AF143668
<i>smez</i> -19	n.f.	-	-	emm44	AF143669
<i>smez</i> -20	SMEZ-20	7.82	0.09	PT2841	AF143670
<i>smez</i> -21	SMEZ-21	6.14	0.3	emm14	AF143671
<i>smez</i> -22	SMEZ-22	6.98	0.3	emm48	AF143672
<i>smez</i> -23	SMEZ-23	-	-	n.d.	AF143673
<i>smez</i> -24	SMEZ-24	8.85	n.d.	ST3018	AF143674

*smez*-8, *smez*-19, and *smez*-23 all have a single base pair deletion causing a frame shift. With the exception of *smez*-3, *smez*-4, and *smez*-8, all analyzed variants could be correlated to a single M/emm type. pI, calculated isoelectric point; P<sub>50</sub>, half-maximal mitogenic potency; n.d., not determined; n.f., nonfunctional (pseudogene).

\*Described previously (reference 28).

<sup>1</sup>Published previously (reference 27).

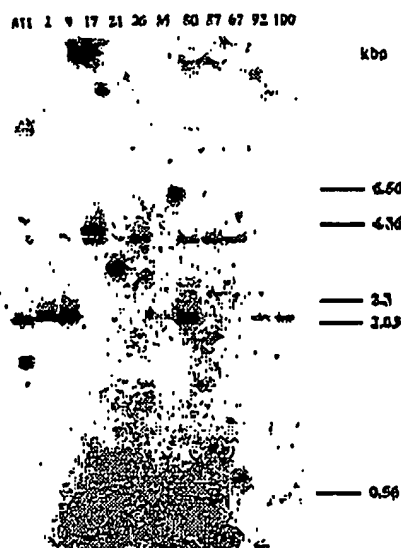


Figure 1. Southern blot analysis of genomic DNA with radiolabeled *smez*. HindIII-digested genomic DNA from various *Streptococcus* isolates was hybridized with a radiolabeled *smez* probe. Band A is a 1953-bp HindIII DNA fragment that carries the *smez* gene. Bands B and C are DNA fragments of ~4 and 4.2 kbp, respectively, which both carry a *smez*-like region. M1, *S. pyogenes* reference strain (M1 type, 700294; American Type Culture Collection); 2, isolate 9639 (MNT); 8, isolate 11789 (MNT); 7, isolate 11152 (PT2612 type); 21, isolate RC4083 (group C *Streptococcus*); 28, isolate 11070 (emm65 type); M, DNA marker lane; 50, isolate 4202 (NZ5118/M82 type); 57, isolate 94/229 (M49 type); 67, isolate 11610 (emm57 type); 82, isolate 95/127 (NZ1437/M89 type); and 100, isolate 94/330 (M4 type).

in one M4 and in one M81 type strain. Four of the five MNT isolates carry *smex-6* and one carries *smex-3*.

With the exception of *smex-6*, *smex-19*, and *smex-23*, which have individual single base pair deletions causing frame shifts, all *smex* alleles possess complete open reading frames. *smex-2*, which was isolated from a separate strain, was not found in any of the current panel of isolates. The 21 intact *smex* genes encode 21 different SMEZ variants.

Sequence identities between *smex* alleles are between 94.5 (*smex-1* and *smex-2*) and 99.9% (equivalent to one base pair difference), e.g., *smex-4* and *smex-7*. A family tree of the 24 *smex* alleles was established based on DNA sequence homology (Fig. 2). The *smex* alleles were grouped into two clusters. Group A comprises *smex-1*, *smex-5*, *smex-8*, *smex-10*, *smex-12*, and *smex-15*. All other *smex* alleles are in group B. Sequence variation between all *smex* alleles is restricted to only 62 nucleotide positions, or 9.8% of the coding region for the mature protein (Table II). Most of the variations are G/A and T/C transitions, while transversions only occur at position 290 (A/C), 291 (A/T), and 358 (G/C). 31 of the variable positions (50%) result in amino acid changes; the other 50% are silent exchanges. The frequency of a nucleotide exchange at a particular po-

sition also varies significantly within the 24 *smex* alleles. 17 of the 62 sites are point mutations for single alleles, while the remaining 45 sites are point mutations common to 2 or more alleles. The most variable sites are positions 231, 234, and 238 and occur in 12 alleles (50%). Some of these highly variable nucleotides are linked together and build five distinct regions (Table III). One region (III) comprising 10 polymorphic sites between nucleotide 282 and 298 occurs in only 2 combinations within the 24 *smex* alleles. These combinations segregate the two main groups of the *smex* family tree. The other four blocks are less conserved, and sequence combinations between these blocks occur in several alleles, generating gene mosaics. Interestingly, the polymorphic sequence blocks are more frequent in the center of the gene, between nucleotide positions 231 and 427. Only one polymorphic block (V) is in the last third of the gene. No blocks occur in the first third. The sequence clusters and mosaic distribution are also reflected in the amino acid sequence (Table IV). Interestingly, the five polymorphic amino acid positions within block III are all between amino acid positions 91 and 100 and are therefore clustered closer together than in the other blocks. The amino acid composition of this polymorphic block influences the isoelectric point of the SMEZ variants. All group A variants have a pI between 5.02 and 5.4, whereas the pIs of group B SMEZ are between 5.73 and 9.62 (Table II).

**Expression of SMEZ Variants by *S. pyogenes* Isolates.** Both SMEZ-1 and SMEZ-2 specifically target human V $\beta$ 8-bearing T cells (27). Jurkat cells (a V $\beta$ 8 T cell line) were used to determine SMEZ expression from *S. pyogenes* culture supernatants from all 103 streptococcal isolates (see Materials and Methods). 62 isolates (66% of the GAS isolates) showed V $\beta$ 8-stimulating activity. No V $\beta$ 8 activity was detected from the nine non-GAS, the four *smex* PCR-negative GAS strains, and the isolates carrying the pseudogenes *smex-6*, *smex-19*, or *smex-23*. These results suggest that SMEZ is the only V $\beta$ 8-targeting SA $\alpha$ g in *S. pyogenes*. Culture supernatant from 12 isolates with complete open reading frames for *smex* did not stimulate Jurkat cells. These isolates possess *smex-5*, *smex-10*, *smex-12*, *smex-18*, *smex-20*, *smex-21*, or *smex-22*. These results suggest the following possibilities: (a) those isolates do not express SMEZ under the selected growth conditions; (b) mutations outside the *smex* coding region repress gene expression; or (c) these isolates express SMEZ variants that have lost their specificity to target V $\beta$ 8 T cells. Recombinant SMEZ variants were generated in *E. coli* using the GST expression system.

Successful recombinant expression was achieved for rSMEZ-3, rSMEZ-5, rSMEZ-7, rSMEZ-8, rSMEZ-9, rSMEZ-13, rSMEZ-20, rSMEZ-21, and rSMEZ-22. Fusion proteins were completely soluble and gave yields of ~30 mg/liter. The purified recombinant toxins were applied to isoelectric focusing (Fig. 3). Isoelectric points between SMEZ variants were consistent with the calculated pIs (Table II).

Native conformation of the recombinant toxins was confirmed by human T cell proliferation assay. All purified SMEZ variants were active on human T cells with similar

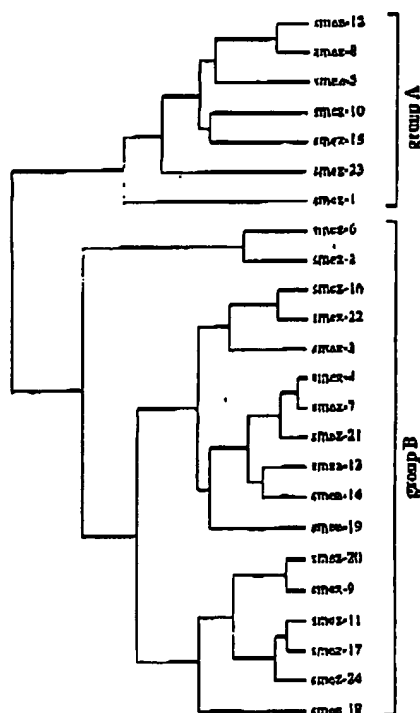


Figure 2. *smex* gene family tree. The family tree was created using the PileUp program v8 (GeneC Computer Group) and is based on nucleotide sequence homology. The *smex* family can be divided into two major groups based on a conserved DNA region (region III). *smex-6*, *smex-19*, and *smex-23* are pseudogenes with a single base pair deletion causing a frame shift.

Table III. DNA Sequence Alignment of *smex* Alleles

	I																	II			III											
Allele	33	49	51	57	68	70	78	131	135	138	151	159	184	185	191	195	207	251	254	258	245	251	252	261	264	267	271	273	276	279	281	
group A																																
smex-1	T	T	A	A	C	C	G	T	T	T	G	T	G	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-5	T	G	G	A	A	C	C	G	T	T	G	T	G	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-8	T	T	G	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-10	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-12	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-15	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-23	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
group B																																
smex-2	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-3	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-4	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-6	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-7	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-9	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-11	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-13	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-14	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-16	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-17	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-18	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-19	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-20	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-21	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-22	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
smex-24	T	T	C	A	A	C	C	G	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A	A	C	C	G	A	A	A	A	A
diversity	0.04	0.02	0.10	0.04	0.08	0.04	0.08	0.04	0.04	0.08	0.04	0.04	0.16	0.31	0.04	0.16	0.16	0.5	0.5	0.5	0.04	0.46	0.04	0.16	0.33	0.41	0.04	0.29	0.29	0.29	0.04	
codon pos	3	1	3	3	2	1	3	2	3	3	1	3	1	2	2	3	1	3	3	1	2	2	3	3	3	3	3	1	3	3	3	2

	III								IV										V													
Allele	282	284	286	289	290	291	292	297	298	311	331	342	343	348	349	351	354	357	358	427	473	475	477	503	514	525	529	530	553	563	582	
group A																																
smex-1	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-3	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-8	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-10	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-12	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-15	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-23	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
group B																																
smex-2	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-3	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-4	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-6	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-7	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-9	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-11	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-13	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-14	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-16	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-17	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-18	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-19	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-20	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-21	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-22	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
smex-24	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	A	T	A	A	A	A	A	A	A	T	T	T	T
diversity	0.29	0.04	0.29	0.30	0.29	0.29	0.29	0.08	0.29	0.04	0.29	0.29	0.29	0.29	0.04	0.29	0.29	0.29	0.29	0.25	0.16	0.16	0.16	0.08	0.38	0.18	0.38	0.38	0.08	0.04	0.04	
codon pos	3	2	1	1	2	3	1	3	1	2	1	3	1	3	1	3	3	3	3	1	1	2	1	3	2	1	3	1	2	3	1	3

Only the 52 variable nucleotides (8.8% of the *smex* gene) are shown, of which 17 are point mutations for single alleles and 45 are common to 2 or more alleles. Some of the variable sites are linked together, building sequence blocks (shaded gray) within five distinct regions (I–V). In region III, only two sequence combinations occur, correlating with the two major branches (groups A and B) of the *smex* family tree (see Fig. 2). Combinations between sequence blocks within the other four regions in different alleles generate gene mosaics.



Table IV. Amino Acid Alignment of SMEZ Alleles

group A

						I								II								VI								V					
Allele	17	23	34	26	44	62	63	68	80	62	84	91	94	96	97	98	99	100	104	111	115	117	120	149	153	158	165	172	177	188	192				
SMEZ-1	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N				
SMEZ-3	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N				
SMEZ-8	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N				
SMEZ-10	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N				
SMEZ-12	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N				
SMEZ-15	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N				

group B

SMEZ-2	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N
SMEZ-3	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N
SMEZ-4	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N
SMEZ-7	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N
SMEZ-9	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N
SMEZ-11	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N
SMEZ-13	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N
SMEZ-14	M	T	L	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N
SMEZ-16	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N
SMEZ-17	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N
SMEZ-18	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N
SMEZ-20	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N
SMEZ-21	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N
SMEZ-22	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N
SMEZ-24	V	T	V	D	R	K	K	K	K	K	K	Y	Y	Y	Y	Y	M	P	P	P	P	P	P	P	P	K	K	D	D	V	N

diversity

0.1	0.05	0.05	0.1	0.05	0.38	0.19	0.19	0.48	0.09	0.43	0.33	0.03	0.29	0.29	0.39	0.1	0.29	0.05	0.33	0.33	0.05	0.33	0.29	0.05	0.14	0.05	0.38	0.38	0.05	0.03
-----	------	------	-----	------	------	------	------	------	------	------	------	------	------	------	------	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------

Only the 31 variable residues are shown (14.9% of the mature SMEZ protein). The sequence blocks and mosaic structure seen in the DNA sequence alignment (Table III) are also reflected in the amino acid sequence.

half-maximal responses ( $P_{50}$  values) between 0.09 and 0.3 pg/ml, except for rSMEZ-5 with a  $P_{50}$  of 0.7 pg/ml, and rSMEZ-10 and rSMEZ-12, which gave  $P_{50}$  values of 2 and 15 pg/ml, respectively (Table II).

rSMEZ-5, rSMEZ-20, rSMEZ-21, and rSMEZ-22 all stimulated Jurkat cells (Fig. 4), thus confirming that the lack of V $\beta$ 8-stimulating activity of wild-type streptococcal culture supernatants is due to lack of expression rather than altered V $\beta$  specificity of the SMEZ variants produced.

**Antigenic Variation in SMEZ Variants.** Human sera from 100 healthy donors were tested for the presence of naturally acquired neutralizing antibodies to SMEZ variants rSMEZ-1, rSMEZ-2, rSMEZ-3, rSMEZ-5, rSMEZ-7, rSMEZ-8, rSMEZ-9, rSMEZ-13, rSMEZ-20, rSMEZ-21,

and rSMEZ-22. The neutralizing response was determined as the percentage of inhibition of T cell stimulation by SMEZ in the presence of 10% human serum versus 10% FCS. More than 75% inhibition was defined as full, 30–75% as moderate, and <30% as negative neutralization.



Figure 3. Isoelectric focusing gel. 2 µg of purified recombinant toxin was run on a 5.5% polyacrylamide isoelectric focusing gel (pH 5–8). The gel shows significant variation of the isoelectric points between SMEZ variants. Toxins with low isoelectric points, such as rSMEZ-5 and rSMEZ-8, belong to group A of the SMEZ family. Members of group B are characterized by a more basic isoelectric point.

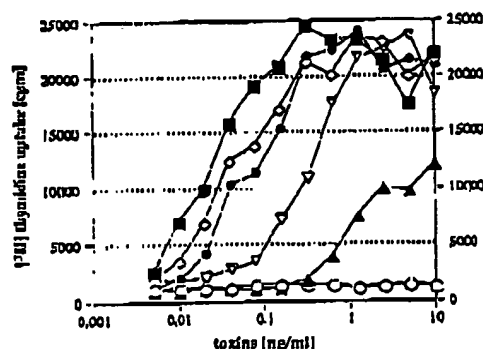


Figure 4. Jurkat cell assay. Jurkat cells (bearing a V $\beta$ 8 TCR) and LO-2 cells were mixed with varying concentrations of recombinant toxin and incubated for 24 h, before Scl cells were added. After 1 d, 0.1 mCi [ $^3$ H]thymidine was added and cells were counted after an additional 24 h. O, unstimulated;  $\nabla$ , rSMEZ-1;  $\Delta$ , rSMEZ-5;  $\blacksquare$ , rSMEZ-20;  $\circ$ , rSMEZ-21; and  $\bullet$ , rSMEZ-22.

Table V. Neutralization Assay

Group A	% inhibition																			
	22	26	28	30	31	32	34	35	36	38	40	58	73	75	86	90	91	99	100	
SMEZ-1	93.0	93.0	87.5	26.9	64.9	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	
SMEZ-6	94.7	94.9	32.8	73.0	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	
SMEZ-8	88.5	14.8	90.5	5.3	46.9	45.9	95.4	95.4	95.4	15.1	86.5	95.4	95.4	95.4	23.6	95.4	95.4	95.4	11.1	

Group B	% inhibition																			
	22	26	28	30	31	32	34	35	36	38	40	58	73	75	86	90	91	99	100	
SMEZ-2	71.7	85.1	93.4	77.2	21.5	3.0	75.1	24.1	59.0	39.8	95.4	95.4	20.5	44.3	22.8	95.4	95.4	4.8	0.6	
SMEZ-3	5.7	99.4	34.0	29.8	18.2	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	13.4	95.4	
SMEZ-7	18.5	98.5	46.4	18.2	95.4	95.4	95.4	95.4	95.4	95.4	95.4	95.4	13.4	87.3	4.7	95.4	95.4	95.4	95.4	
SMEZ-4	3.8	90.5	27.9	44.2	64.8	95.4	95.4	95.4	95.4	95.4	95.4	95.4	45.0	71.3	28.1	95.4	95.4	95.4	95.4	
SMEZ-13	18.0	89.5	73.3	34.9	33.5	95.4	95.4	95.4	95.4	95.4	95.4	95.4	11.8	72.3	2.4	95.4	95.4	11.5	95.4	
SMEZ-20	3.8	95.9	1.1	49.0	48.5	95.4	95.4	95.4	95.4	11.4	95.4	95.4	15.5	49.5	7.3	95.4	95.4	16.0	95.4	
SMEZ-21	6.1	89.5	22.5	34.9	47.5	95.4	95.4	95.4	95.4	2.5	95.4	95.4	3.3	64.7	2.0	95.4	95.4	18.6	16.1	
SMEZ-22	5.9	95.9	21.5	25.0	62.8	95.4	95.4	95.4	95.4	4.1	95.4	95.4	6.3	78.3	7.3	95.4	95.4	3.8	28.8	

Human sera from 100 healthy New Zealand donors were used to study the presence of antibodies neutralizing the mitogenic activity of recombinant SMEZ variants. The figures given are the percentage of inhibition compared with a reference test in which human serum was replaced by FCS. The table shows the results obtained with 20 selected sera. The remaining 80 sera all showed high neutralizing responses for the whole panel of variants.

85 sera fully neutralized the activity of all 11 variants tested. The remaining 15 sera displayed only partial neutralizing activity on a subset of the variants. 20 sera were selected for a repeat test, and the confirmed results are shown in Table V. Seven sera (26, 32, 39, 73, 86, 99, and 100) displayed very different neutralizing response profiles, suggesting the existence of variant (epitope) specific antibodies. For example, serum 86 strongly neutralized rSMEZ-1 and rSMEZ-5, but none of the other variants. Interestingly, both SMEZ-1 and SMEZ-5 belong to group A and share a common amino acid stretch in region III (Table IV). However, serum 86 did not effectively neutralize another group A variant, rSMEZ-8. The only amino acid differences between SMEZ-1, SMEZ-5, and SMEZ-8 that could explain this different neutralization are residues D62A, M68V (both in region I), and P117T (in region VI). Thus, for serum 86, the likely neutralizing epitopes must include one of these three residues.

Serum 39 was similar to serum 86, and sera 73 and 75 also neutralized rSMEZ-8-stimulated T cells. SMEZ-1 and SMEZ-5 were commonly neutralized by most of the sera tested. Only two sera (30 and 31) displayed less than complete neutralization responses. In contrast, SMEZ-2 activity was not inhibited by eight sera, and another six sera showed only moderate inhibition. SMEZ-2 was therefore the least commonly neutralized variant.

## Discussion

SMEZ is the most potent bacterial SAg so far discovered, with specificity towards V $\beta$ 4- and V $\beta$ 8-bearing human T

cells. Unlike all other SAgS, SMEZ is highly polymorphic, having arisen from genetic mosaicism at a single genetic locus. 22 different *smcz* alleles were found in 21 different M-type positive GAS isolates obtained over a 15-yr period throughout New Zealand. In addition to the 2 alleles, *smcz* and *smcz-2*, described previously (27, 28). All of the 94 isolates tested positive for the *smcz* gene by PCR or Southern analysis. Thus, *smcz* is not only the most potent SAg, it is also the most prevalent. The polymorphism in the SMEZ protein results in antigenic variants that are not consistently neutralized by naturally occurring antibodies present in normal adult human sera, yet the human TCR  $\beta$  chain profile remains steadfastly V $\beta$ 4 and V $\beta$ 8.

**Genetic Mosaicism.** The variation in *smcz* has most likely arisen through gene mosaicism, and the strongest indicator for this is that 30 out of 31 polymorphic positions have only 2 possible amino acids. The mechanism for gene mosaicism is homologous recombination between two copies of a gene, and Southern blot analysis shows that some isolates indeed carry a second *smcz* gene or *smcz*-like sequence. However, homologous recombination of *smcz-1* with the *smcz*-like sequence on the M1 genome would result in another pseudogene due to a nonsense mutation within the *smcz*-like sequence. Moreover, >2 homologous genes would be needed to generate the mosaic pattern seen in the 24 *smcz* alleles. The different sequence segments appear to have originated from different strains, suggesting the possibility of intergenic recombination with a horizontally transferred homologous gene from another strain. Inter-strain transfer followed by homologous recombination has

been reported to play a key role in the evolution of *emm*-like genes (30), indicating that such a mechanism exists in *S. pyogenes*.

Sequence combination between different alleles is not the only mechanism to create variability in the *smez* gene. 11 of the 31 variable residues of the SMEZ variants correlate with individual point mutations for single alleles and must have occurred independently in the respective strains. This is also remarkable, given the fact that such a high mutation rate is not observed in any other *sag* genes, including *spe-g*, which was also found in all isolates tested. It is possible that point mutations in other *sag* genes are eliminated by selective pressure in order to maintain functionality of the toxin. However, this seems unlikely, since the variant residues in SMEZ do not contribute to the overall structure and are not conserved within the SAg family.

Allelic variation has been described for the staphylococcal enterotoxin C (SEC). SEC1 and SEC2 differ by only seven amino acid residues, but target a different set of TCR (31). Also, SEA and SEE, which still share 85% amino acid sequence identity, vary in their TCR specificity (32).

Protein expression of a V $\beta$ 8-stimulating toxin was observed in the culture supernatant of 66% of the GAS isolates. No V $\beta$ 8 activity was found in any of the isolates that carry a defective *smez* gene, suggesting that *smez* is the only V $\beta$ 8-specific toxin in *S. pyogenes*. Since the *smez* gene has been analyzed in only a fraction of the GAS panel, it is possible that more nonfunctional *smez* alleles occur in the natural *S. pyogenes* population. Of the identified complete *smez* genes, 7 out of 21 did not express the SMEZ protein, suggesting a tight control of gene transcription. However, it is important not to assume that expression under *in vitro* conditions always equates to expression *in vivo*. Interestingly, the SMEZ-producing strains include M1, M4, and M12 types. These are connected to STLS (33), whereas the SMEZ-13-producing NZ1437/M89 type strain is connected to rheumatic fever in New Zealand (34). A larger panel of GAS isolates, in particular those known to cause invasive diseases, must be analyzed for SMEZ expression to identify any connection between expression of a certain SMEZ variant and streptococcal diseases.

**Antigenic Variation.** There are 31 variable amino acid positions in SMEZ. Analysis of a recent SMEZ-2 crystal structure shows that 26 of these residues are surface exposed and scattered across the entire protein. Only five of the variable positions have side chains that are buried, and all are conservative changes, indicating that there has been selective pressure to vary the surface of the molecule while preserving structural integrity of the molecule (34a). Widespread prevalence of SMEZ-producing strains of *S. pyogenes* in the normal adult human population was confirmed by seroconversion in 100/100 normal healthy young adults. However, 15% of the sera showed only partial neutralization profiles against a panel of recombinant SMEZ variants. SMEZ-1 was the most commonly neutralized, whereas SMEZ-2 was the least. SMEZ-2 is a rare variant that has not yet been found in any *S. pyogenes* isolate in New Zealand. SMEZ-2 was identified in a single isolate of un-

known M/*emm* type from a patient with rheumatic fever in the United States, and is the most potent variant known to date (27). In contrast, SMEZ-1 was found associated with the M1 strain, a very common strain in New Zealand. These data suggest that certain SMEZ variants are better capable of generating a generic protective neutralizing response against all SMEZ variants than others, and that protective antibodies offer only partial protection against a subset of SMEZ variants. This raises an interesting question of whether individual SMEZ variants are linked to disease severity. Screening patients with severe streptococcal-mediated disease for lack of seroconversion against individual SMEZ variants might provide information in this regard.

The strong linkage between the *smez* gene and the *emm* gene, which codes for the major polymorphic surface antigen M protein, offers one explanation for the mechanisms of *smez* genetic variation. The *smez* gene is located 14 kb upstream from the *Vir* regulon. The M protein represents the major neutralizing surface antigen on *Streptococci*, and generation of diversity in the M protein and antigenic determinants might also affect neighboring genes. Thus far, we have analyzed only 21 different M/*emm* types for the *smez* gene. It is likely that the actual number of *smez* alleles among the *S. pyogenes* population is much higher. This might be of particular interest, as the epidemiology of streptococcal disease appears to be connected to certain M types, and temporal shifts in prevalence of certain M types occur (35).

**Role of SMEZ.** SAgS interfere with host adaptive immune responses, driving profound Th1-type responses in an antigen-nonspecific way characterized by high levels of INF- $\gamma$ , IL-2, and nonspecific T cell proliferation. How this response improves survival of the bacteria has yet to be clearly established, but one might argue that the prime function is to prevent the production of high affinity cytotoxic antibodies by promoting a cellular, Th1-mediated response. The high degree of antigenic variation is good evidence that SMEZ provides a significant growth advantage for *S. pyogenes* and that the variation at this locus has been driven primarily by the need to escape antibody neutralization, rather than through a need to expand its T cell repertoire.

**Retention of the TCR Specificity.** In contrast to the antigenic diversity, the TCR V $\beta$  binding site of SMEZ appears to be unaffected by surrounding amino acid variation. One conclusion to draw from this is that V $\beta$ 4- and V $\beta$ 8-bearing T cell subsets are uniquely important in host defence against *Streptococci*. However, one could also argue that retention of this T cell specificity is due to the physical constraints of the SMEZ-TCR  $\beta$  chain interaction, so that other TCR  $\beta$  chains cannot be effectively accommodated without loss of SAg activity. However, the fact that SMEZ effectively targets murine TCR  $\beta$  chains (27) suggests that there are no special constraints on the SMEZ-TCR binding site. Why SMEZ should be privy to such extensive variation while all other SAgS are substantially invariant is unclear, but does point to an important role of this particular SAg in the survival of all group A *Streptococci*. There are

examples of allelic variation in other streptococcal SAGs, such as SPEA (24), SPE-C (36), and SSA (37), but to a much lesser degree and with all the mosaicism seen at the *speA* locus. Genes with mosaic structures are known in other pathogenic organisms such as streptokinase (*skn* [38]), the *emm*-like genes (30), or the *Streptococcus pneumoniae* autolysin gene (*lytA* [39]), and each codes for an important virulence factor.

The antigenic variation appears to have arisen through the selective advantage that SAG escape mutants provide to the species *S. pyogenes*. This is perhaps the best evidence yet that the SAGs confer an advantage on the strain that produces it. Moreover, the ubiquitous expression and genetic variation tend to suggest that *speA*, more than other SAGs, significantly enhances the survival of group A *Streptococcus* in general and thus might be linked to increased virulence and streptococcal disease severity. The *speA* locus may provide an important genotyping locus to examine virulence and disease susceptibility, and may also offer a potential target for streptococcal vaccine development.

We wish to thank Dr. Lachy McLean for supplying us with human sera.

This study was supported by the Health Research Council of New Zealand.

Submitted: 8 December 1999

Revised: 6 March 2000

Accepted: 9 March 2000

## References

1. Bronze, M.S., and J.B. Dale. 1986. The reemergence of serious group A streptococcal infections and acute rheumatic fever. *Am. J. Med. Sci.* 311:41-54.
2. Demers, B., A.B. Simor, H. Vellend, P.M. Schlievert, S. Byrne, F. Jamieson, S. Walmsley, and D.E. Low. 1993. Severe invasive group A streptococcal infections in Ontario, Canada: 1987-1991. *Clin. Infect. Dis.* 18:792-800.
3. Hoge, C.W., B. Schwartz, D.F. Talkington, R.F. Brelman, E.M. MacNeill, and S.J. Engender. 1993. The changing epidemiology of invasive group A streptococcal infections and the emergence of streptococcal toxic shock-like syndrome. A retrospective population-based study. *JAMA (J. Am. Med. Assoc.)* 269:384-389.
4. Low, D.E., B. Schwartz, and A. McGeer. 1997. The reemergence of severe group A streptococcal disease: an evolutionary perspective. In *Emerging Infections*. W.M. Scheld, D. Armstrong, and J.M. Hughes, editors. American Society for Microbiology, Washington, D.C. 93-123.
5. Holm, S.E., A. Norrby, A.M. Bergholm, and M. Norgren. 1992. Aspects of pathogenesis of serious group A streptococcal infections in Sweden, 1988-1989. *J. Infect. Dis.* 166: 31-37.
6. Bianco, A.L. 1991. Group A streptococcal infections and acute rheumatic fever. *N. Engl. J. Med.* 325:783-793.
7. Cone, L.A., D.R. Woodard, P.M. Schlievert, and G.S. Tomory. 1987. Clinical and bacteriologic observations of a toxic shock-like syndrome due to *Streptococcus pyogenes*. *N. Engl. J. Med.* 317:146-149.
8. Stevens, D.L., M.H. Tanner, J. Winship, R. Swartz, K.M. Ries, P.M. Schlievert, and E. Kaplan. 1989. Severe group A streptococcal infections associated with a toxic shock-like syndrome and scarlet fever toxin A. *N. Engl. J. Med.* 321:1-7.
9. Forni, A.L., E.L. Kaplan, P.M. Schlievert, and R.B. Roberts. 1995. Clinical and microbiological characteristics of severe group A streptococcus infections and streptococcal toxic shock syndrome. *Clin. Infect. Dis.* 21:333-340.
10. Alouf, J.E., H. Knoch, and W. Koehler. 1991. The family of mitogenic, shock-inducing and superantigenic toxins from staphylococci and streptococci. In *The Sourcebook of Bacterial Protein Toxins*, J.E. Alouf and J.H. Freer, editors. Academic Press, San Diego. 367-414.
11. Marrack, P., and J. Kappler. 1990. The staphylococcal enterotoxins and their relatives. *Science* 248:705-711.
12. Past, D.J., P.M. Schlievert, and R.D. Nelson. 1989. Toxic shock syndrome-associated staphylococcal and streptococcal pyrogenic toxins are potent inducers of tumor necrosis factor production. *Infect. Immun.* 57:291-294.
13. Herman, A., J.W. Kappler, P. Marrack, and A.M. Pullen. 1991. Superantigens: mechanism of T-cell stimulation and role in immune responses. *Annu. Rev. Immunol.* 9:745-772.
14. Kozin, B.L., D.Y. Leung, J. Kappler, and P. Marrack. 1993. Superantigens and their potential role in human disease. *Adv. Immunol.* 54:99-166.
15. Abe, J., B.L. Kozin, K. Jujo, M.E. Melish, M.P. Glode, T. Kohsaka, and D.Y. Leung. 1992. Selective expansion of T cells expressing T-cell receptor variable regions V $\beta$ 2 and V $\beta$ 8 in Kawasaki disease. *Proc. Natl. Acad. Sci. USA* 89: 4066-4070.
16. Leung, D.Y., R.C. Giorno, L.V. Kazemi, P.A. Flynn, and J.B. Busse. 1995. Evidence for superantigen involvement in cardiovascular injury due to Kawasaki syndrome. *J. Immunol.* 155:5018-5021.
17. Talkington, D.F., B. Schwartz, C.M. Black, J.K. Todd, J. Elliott, R.F. Brelman, and R.R. Facklam. 1993. Association of phenotypic and genotypic characteristics of invasive *Streptococcus pyogenes* isolates with clinical components of streptococcal toxic shock syndrome. *Infect. Immun.* 61:3369-3374.
18. Musser, J.M., A.R. Hauser, M.H. Kim, P.M. Schlievert, K. Nelson, and R.K. Selander. 1991. *Streptococcus pyogenes* causing toxic-shock-like syndrome and other invasive diseases: clonal diversity and pyrogenic exotoxin expression. *Proc. Natl. Acad. Sci. USA* 88:2668-2672.
19. Hackett, S.P., and D.L. Stevens. 1992. Streptococcal toxic shock syndrome: synthesis of tumor necrosis factor and interleukin-1 by monocytes stimulated with pyrogenic exotoxin A and streptolysin O. *J. Infect. Dis.* 165:879-885.
20. Hauser, A.R., D.L. Stevens, E.L. Kaplan, and P.M. Schlievert. 1991. Molecular analysis of pyrogenic exotoxins from *Streptococcus pyogenes* isolates associated with toxic shock-like syndrome. *J. Clin. Microbiol.* 29:1562-1567.
21. Yu, C.E., and J.J. Ferretti. 1989. Molecular epidemiologic analysis of the type A streptococcal exotoxin (erythrogenic toxin) gene (*speA*) in clinical *Streptococcus pyogenes* strains. *Infect. Immun.* 57:3715-3719.
22. Chaussee, M.S., J. Liu, D.L. Stevens, and J.J. Ferretti. 1996. Genetic and phenotypic diversity among isolates of *Streptococcus pyogenes* from invasive infections. *J. Infect. Dis.* 173:901-908.
23. Haueh, P.R., J.J. Wu, P.J. Tsai, J.W. Liu, Y.C. Chuang, and K.T. Luh. 1998. Invasive group A streptococcal disease in Taiwan is not associated with the presence of streptococcal pyrogenic exotoxin genes. *Clin. Infect. Dis.* 26:584-589.

24. Nelson, K., P.M. Schlievert, R.K. Selander, and J.M. Musser. 1991. Characterization and clonal distribution of four alleles of the *speA* gene encoding pyrogenic exotoxin A (scarlet fever toxin) in *Streptococcus pyogenes*. *J. Exp. Med.* 174:1271-1274.
25. Musser, J.M., V. Kapur, S. Kanjilal, U. Shah, D.M. Musher, N.L. Barg, K.H. Johnston, P.M. Schlievert, J. Henrichsen, D. Garlach, et al. 1993. Geographic and temporal distribution and molecular characterization of two highly pathogenic clones of *Streptococcus pyogenes* expressing allelic variants of pyrogenic exotoxin A (Scarlet fever toxin). *J. Infect. Dis.* 167:337-346.
26. Kline, J.B., and C.M. Collins. 1996. Analysis of the superantigenic activity of mutant and allelic forms of streptococcal pyrogenic exotoxin A. *Infect. Immun.* 64:861-869.
27. Proft, T., S.L. Moffatt, C.J. Berkahn, and J.D. Fraser. 1999. Identification and characterization of novel superantigens from *Streptococcus pyogenes*. *J. Exp. Med.* 189:89-102.
28. Kamezawa, Y., T. Nakahara, S. Nakano, Y. Abe, J. Nozaki-Renard, and T. Isono. 1997. Streptococcal mitogenic exotoxin Z, a novel acidic superantigenic toxin produced by a T1 strain of *Streptococcus pyogenes*. *Infect. Immun.* 65:3828-3833.
29. Maniatis, T., E.F. Fritsch, and J. Sambrook. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
30. Whatmore, A.M., and M.A. Kehoe. 1994. Horizontal gene transfer in the evolution of group A streptococcal *emm-ilk* genes: gene mosaics and variation in *Vir* regulons. *Mol. Microbiol.* 11:363-374.
31. Deringer, J.R., R.J. Ely, C.V. Stauffacher, and G.A. Bohach. 1996. Subtype-specific interactions of type C staphylococcal enterotoxins with the T-cell receptor. *Mol. Microbiol.* 22:523-534.
32. Hudson, K.R., H. Robinson, and J.D. Fraser. 1993. Two adjacent residues in staphylococcal enterotoxins A and E determine T cell receptor V $\beta$  specificity. *J. Exp. Med.* 177:175-184.
33. Zurawski, C.A., M. Bardsley, B. Beall, J.A. Elliott, R. Packiam, B. Schwartz, and M.M. Farley. 1998. Invasive group A streptococcal disease in metropolitan Atlanta: a population-based assessment. *Clin. Infect. Dis.* 27:160-167.
34. Martin, D.R., L.M. Voss, S.J. Walker, and D. Lennon. 1994. Acute rheumatic fever in Auckland, New Zealand: spectrum of associated group A streptococci different from expected. *Ped. Infect. Dis. J.* 13:264-269.
- 34a. Arcus, V., T. Proft, J.A. Sirelli, H.M. Baker, J.D. Fraser, and E.N. Baker. 2000. Conservation and variation in superantigen structure and activity highlighted by 3-dimensional structures of two new superantigens from *Streptococcus pyogenes*. *J. Mol. Biol.* In press.
35. Martin, D., J. Kakani, and J. Szeto. 1997. Clonal analysis of five M types causing most disease in New Zealand. *Adv. Exp. Med. Biol.* 418:331-333.
36. Kapur, V., K. Nelson, P.M. Schlievert, R.K. Selander, and J.M. Musser. 1992. Molecular population genetic evidence of horizontal spread of two alleles of the pyrogenic exotoxin C gene (*speC*) among pathogenic clones of *Streptococcus pyogenes*. *Infect. Immun.* 60:3513-3517.
37. Reda, K.B., V. Kapur, D. Goels, J.G. Lamphear, J.M. Musser, and R.R. Rich. 1996. Phylogenetic distribution of streptococcal superantigen SSA allelic variants provides evidence for horizontal transfer of *ssa* within *Streptococcus pyogenes*. *Infect. Immun.* 64:1161-1165.
38. Kapur, V., S. Kanjilal, M.R. Hamrick, L.L. Li, T.S. Whitlam, S.A. Sawyer, and J.M. Musser. 1995. Molecular population genetic analysis of the streptokinase gene of *Streptococcus pyogenes*: mosaic alleles generated by recombination. *Mol. Microbiol.* 16:509-519.
39. Whatmore, A.M., and C.G. Dowson. 1999. The autolysin-encoding gene (*lytA*) of *Streptococcus pneumoniae* displays restricted allelic variation despite localized recombination events with genes of pneumococcal bacteriophage encoding cell wall lytic enzymes. *Infect. Immun.* 67:4551-4556.